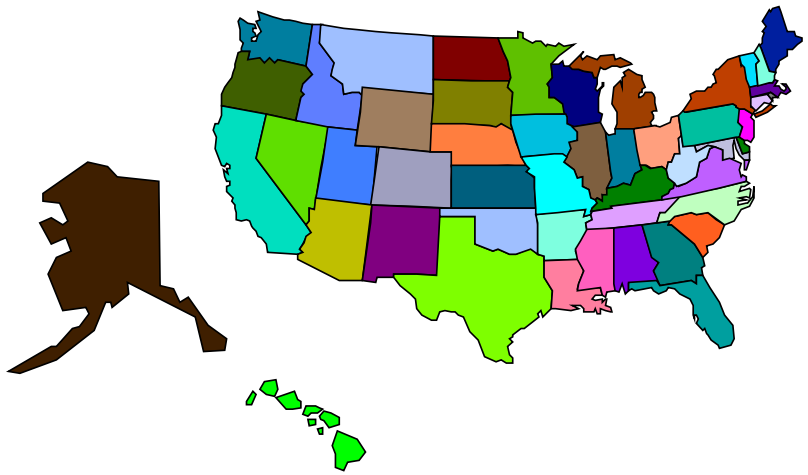
Multimedia/Multimodal Interfaces

CIC Human Centered Systems Working Group (HuCS)

Gary W. Strong

A National Need

- Participation in society is increasingly information-dependent.
  - Competition, mentoring, production, distribution, criticism, and other social activities are information-intensive.
- Multimedia/multimodal interfaces
  - help deter information caste formation;
  - motivate educational goals;
  - promote nomadic computing; and
  - allow citizens with disabilities to lead more independent lives.



The “Media” Issue

- Digital knowledge is rapidly shared in a variety of media (text, video, speech) and languages.
- People have differential modality-dependencies (blind, deaf, physical disabilities) as well as differential literacy.
- Media/language-independent services would:
  - assist access by people with disabilities;
  - assist translation to spoken language for the illiterate;
  - assist machine translation and summarization;
  - guide topic identification/vigilance applications; and
  - promote new device markets.

Status of Research

- HLR was a FY96 $5M NSF/DARPA infrastructure program to create new media research resources.
- STIMULATE is a FY97 $10M NSF/NSA/DARPA/CIA fundamental research program in:
  - New single and multimodal processing algorithms;
  - Basic theories of collaboration and dialogue.

Some Accomplishments

- Speech-to-text and text-to-speech algorithms.
- Rudimentary haptic interfaces.
- Animated agents capable of coordinated motion and lip-synched speech.
- Restricted text <--> figure translations.

Studies and Workshops

- NAS study ("More than Screen Deep") recommends new research to achieve "every-citizen access:"
  - Modality- and medium-independent form for representation of information;
  - Natural, two-way, learnable, and secure interactions with information technology.
- NSF Human-Centered Systems Workshop recommends new research ...
  - that is collaborative around central themes;
  - that involves testbeds and competitions for evaluation;
  - that creates better understanding of metrics and evaluation; and
  - that capitalizes on high-speed networking and digital multimedia.

Other Calls for New Research

- How do different modalities alter expressions of knowledge?
- How do individuals recruit alternative modalities to compensate for limitations in interaction?
- How do individuals produce & receive the same knowledge in alternative forms?
- How can systems transform knowledge into alternative forms?
- Can a common, transmedia representation facilitate the transformation of knowledge across modalities?

Existing State of the Art

- InterLingua for MT;
- Ontologies, WordNet, FrameNet;
- Lexical Semantics for Figure/Text Translation;
- SGML for rendition-independence; and
- Algorithms for topic identification and summarization.

Software/Communication Parallels

Software		Communication
• Object classes	⟷	• Knowledge Representation
• Platform independence	⟷	• Display independence
• Network as computer	⟷	• Delivery service protocols

Possible New Approaches

- Personal plug & play input/output devices, or profiles;
- Protocols for negotiation between clients and servers for modality, language, or bandwidth of service;
- Display-independent & modality-independent information services; and
- “InterMedia”

InterMedia

- A form of knowledge representation for knowledge integration.
- Relates to multiple modalities.
- Operates in semantically-tractable domains.
- Tied to realistic multi-modal data resources.
- Coupled with modality translation algorithms.
- 10-year vision.

Potential Applications (1)

- Access by the blind using a multi-modal interface whereby sound is used to provide peripheral sensation and touch & speech are used to provide local information.
- Virtual, mobile desktops to allow access to one's familiar and structured information environment wherever they may be and whenever they need it.

Potential Applications (2)

- Signing avatars to serve as portable, laptop signing interpreters that understand speech and produce ASL.
- Knowledge support for better understanding of speech accompanied by gesture and other multimodal understanding.
- Both language and media translation using same technologies - GII access support.

Questions to the committee

- Does the NGI need to demonstrate success in M4I's?
- Are M4I's necessary in a learning environment of fully-connected schools?
- Will M4I's "take up" as rapidly as did digital libraries?